ADMINISTRATIVE - INTERNAL USE ONLY

November 1982

DISCUSSION PAPER

Subject: Performance Appraisal - Should we use ratings and, if so, what kind and how many?

The question raised is one that is invariably asked when individuals are all rated at about the same level. An alternate question is, why use ratings when they do not differentiate among those rated? A rating system is used generally with the expectation that some rated elements will appear at either end of the scale and that the remainder will be scattered throughout the middle.

A. Background

1. Since the Agency's inception, management has recognized the value of, and the need for, performance appraisal and has endorsed the concept by including it as an integral part of the personnel management system. Through the years, in attempting to develop a satisfactory system to suit all needs, the format, style, and/or parameters have been changed to some degree no less than twelve times since the first formal report was initiated in 1947. Beginning with a checklist of 17 rating factors at six levels which was not shown to the employee, we have worked through several variations:

-- 20 written elements with employee reassignment preferences, with rater comments concerning good and bad aspects of performance, and with the report being discussed with the employee (1952);

-- a 50-item checklist including ratings on employees' attitudes towards the Agency, their potential and suitability, with the employee not shown the report (1954);

-- a two-part report with the first part rating the employee on performance, which the employee was shown, and with the second part evaluating potential which the employee was not shown (1955);

-- a single sheet report with seven rating levels (1958);

-- a single sheet report with five redefined rating levels (1962);

-- a three-page report with relabeled, redefined rating levels expanded to seven, with introduction of the Advance Work Plan and the Evaluation of Potential as separate forms (1979);

-- return to single sheet format with elimination of the EOP, and a redesign of the AWP (1981).

In reviewing the variations in their times, it appears that the structure of the performance appraisal program often was directly related to the question of whether or not the contents of the evaluation would be shown the

ADMINISTRATIVE - INTERNAL USE ONLY

employee. When it was finally determined that employees had the right to see the material which had an effect on their careers, the problem of designing an effective Performance Appraisal Report (PAR) was made even more difficult.

2. Through the years, we also have struggled with the "rating creep" that seems to be an inevitable occurrence in a performance appraisal system. This problem is not peculiar to the Agency as the Federal establishment at large, particularly agencies involved in merit pay programs, continues to struggle to find a remedy. The only time-honored "cure" seems to be a radical revision which substitutes a new and different rating scale for the one creating the rating creep problem. Unfortunately, any success in stemming "rating creep" is short-lived as raters, reviewers, and employees soon adjust to the new levels and the rating creep begins its almost guaranteed climb.

3. In 1979, we introduced a newly-revised performance appraisal program which included a three-page PAR form, an Advance Work Plan (AWP) and an Evaluation of Potential (EOP). The rating levels were expanded from five to seven, primarily to allow greater differentiation among employees, but partly to reduce the "rating creep" which had over 80 percent of the Agency employees rated "strong" or "outstanding" on the predecessor scale. Shortly after introduction in 1979, the Office of Personnel initiated an evaluation to determine how effectively the program was implemented and how employees and management officials felt about performance appraisal in the Agency generally. (It should be noted that the task force responsible for developing the revised PAR had representation from all parts of the Agency and researched the subject for many months.)

4. In some respects, the new rating scale hardly had a chance to succeed since some management officials immediately published an equivalency chart comparing the new system with the old. As could be expected, the evaluation revealed that the rating levels of the new PAR were comparable to those of the previously used fitness report, and remain so today (attachment A). It can be concluded that, as currently used, the existing Agency PAR system is no more effective than its predecessors. During the evaluation process, employees expressed very strong sentiments against continuous "tinkering" with the PAR, believing that it was important to allow the system to work. Unfortunately, those same employees had little confidence that any improvement would be achieved.

B. Discussion

1. The primary purpose of a performance appraisal report is to provide a formal record of the judgments Agency managers and supervisory officials periodically make concerning the work performance of subordinates. This information is required for making decisions affecting the training, assignment, advancement and retention of employees.

2. The linkage of performance appraisals to job-related criteria is fundamental to a successful appraisal program. Recent court decisions have made it very clear that selection or promotion practices based solely on the

subjective judgments of supervisors will not be accepted. Factors such as initiative, leadership, mobility, flexibility and dedication may be very important to the total job performance but they are very difficult to define and to measure. The use of personality, style, or character factors rather than results obtained in the performance of specific tasks no doubt is one of the reasons why employees criticize the performance appraisal system as being entirely too subjective. Our evaluation revealed that many employees believed that performance appraisals are frequently based on personality considerations, friendships, etc.

3. Our current emphasis has been to press for establishing performance standards at the "fully satisfactory" level (PAR rating level 4). Unfortunately, few people in the Agency -- and according to the literature, in the U.S. -- find a rating level of fully satisfactory to be acceptable. Employees feel they must be above average, excellent, or outstanding even if they essentially do no more than what is expected on the job. In our own case, the higher the grade level, the more difficult it becomes to hold to a level 4 standard because it is popularly assumed that more is expected of a person at the higher levels. This assumption automatically leads to higher ratings and the contribution to rating creep. The counter argument to that assumption is simply that the expectations, i.e., the level 4 standard, for satisfactory performance at higher grades should be higher (tougher) than for lower-graded personnel.

C. Conclusion

1. Research studies support the notion that a rating system grounded on the development of explicit work standards is the preferred approach to performance appraisal. While the number of rating levels may be debated, a minimum of one standard must ultimately be set, i.e., that standard which establishes as objectively as possible what is expected to be done on the job (fully satisfactory performance level). From this one standard, it can be determined that the employee has either failed to satisfy the demands of the job (unsatisfactory) or that the employee has done more than was expected (exceptional).

2. There are three options that might be considered in dealing with the Agency's performance appraisal program. The pros and cons of each are as follows:

a. Reinforce the present system by requiring supervisory personnel to establish explicit standards at level 4 (fully satisfactory) for all grade levels.

PROs:

-- will communicate to employees that management will not abandon a new system when problems are encountered, but will persevere.

-- will not entail major reorientation efforts or training costs.

-- will not jolt employees by any sudden or radical change in management approach to performance appraisal.

-- will afford us time to test the use of performance standards and to experiment with other performance appraisal approaches.

CONs:

-- will perpetuate a system which frustrates the personnel evaluation process by requiring managers to make decisions based on nondiscriminatory ratings.

-- will be difficult to effect a significantly lower average rating level in the Agency since employees are now accustomed to receiving above-average performance ratings.

b. Implement an appraisal system that eliminates ratings.

PROs:

-- provides maximum latitude for describing an employee's work performance.

-- treats employees as individuals and better characterizes the subordinates' strengths, weaknesses, and perceived potential for advancement.

-- permits requiring the evaluator to address specific issues or to answer standard questions posed by employee evaluation panels.

-- avoids using rating labels whose meaning is illusive or variable among rating officials.

CONs:

-- gives the advantage to supervisors who are adept at writing.

-- permits writers to deal in vague generalities having no substance or utility.

-- fails to recognize supervisory reluctance to criticize or identify subordinates' weaknesses in writing.

-- makes comparative evaluation of employees as required by Agency policy difficult because of variance in writing skills among those preparing the performance appraisal.

c. <u>Establish a combination of narrative and a three-level rating system.</u>

PROs:

-- takes advantage of the good features of a rating system using a performance standard coupled with supporting narrative comment.

-- limits the need for performance standards to one.

-- reduces the number of rating levels to the minimum required for basing personnel actions on job performance.

-- recognizes that relatively fine distinctions using multiple-rating levels cannot be made reliably.

CONs:

-- is impractical to use as the primary tool for ranking employees in the comparative evaluation process.

-- would have minimum utility without the application of a performance standard.

D. <u>Recommendations</u>

Since the most promising method of reliably measuring work performance appears to be through the use of performance standards, it is recommended that:

1. our current PAR policy be reaffirmed for the present; that PAR rating level 4 be established as the level at which a fully satisfactory performance standard will be written for all grades; and

2. that OP continue to monitor the performance appraisal program and design an experimental model combining a three-level rating system with provisions for narrative comment. This experimental model should be developed and used on a trial basis during calendar year 1983.

AVERAGE PERFORMANCE RATING BY DIRECTORATE.

AVE

	1979	1982
DCI	5.56	5.73
DDA	5.34	5.43
DDI	5.12	5.20
DDO	5.50	5.64
DST	5.51	5.63

1979 1982   1979 1982   1979 1982   1979 1982   1979 1982   YR

DCI   DDA   DDI   DDO   DST   DIR
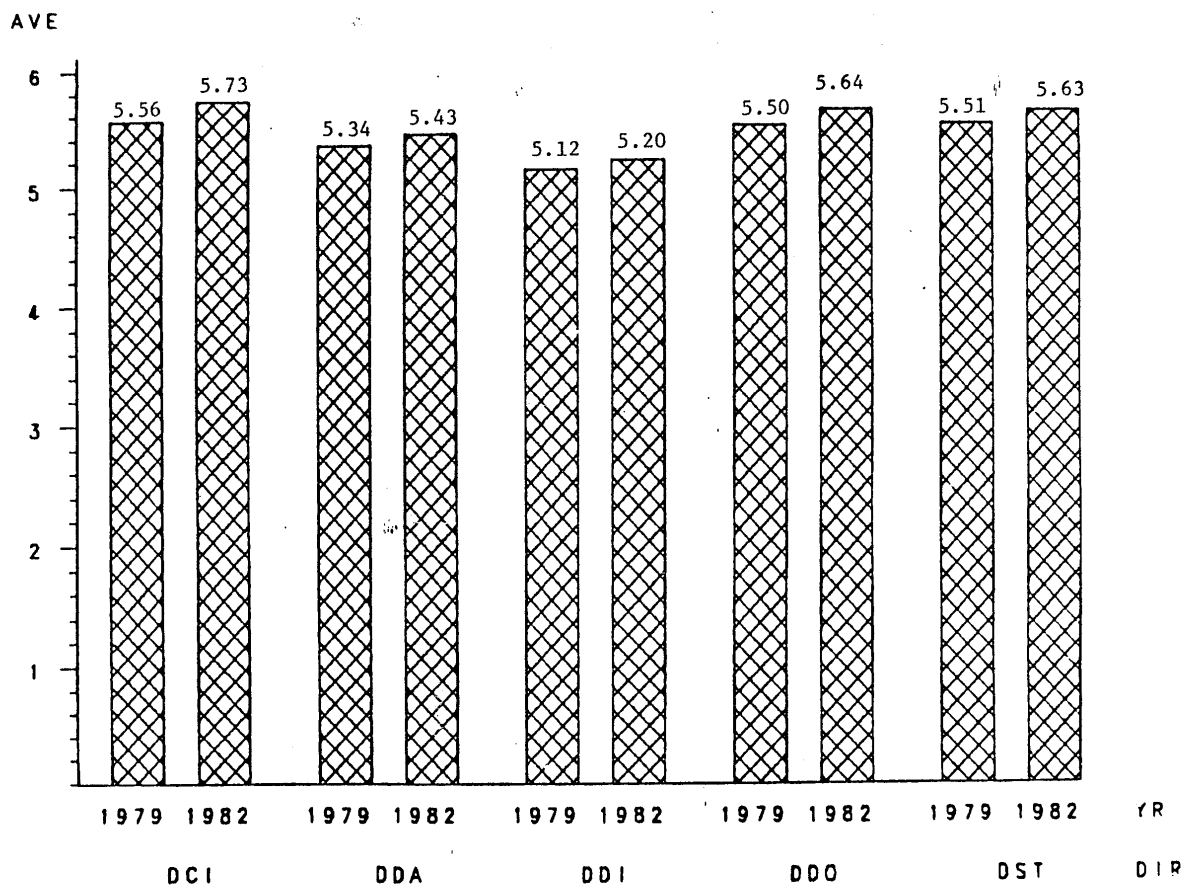
PERFORMANCE RATING FREQUENCY
BY FISCAL YEAR

PCNT

```
80

60

40

20
```

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	
7	8	8	7	8	8	7	8	8	7	8	8	7	8	8	YR
9	0	2	9	0	2	9	0	2	9	0	2	9	0	2	

├─ 1 ─┤	├─ 2 ─┤	├─ 3 ─┤	├─ 4 ─┤	├─ 5 ─┤	SCORE
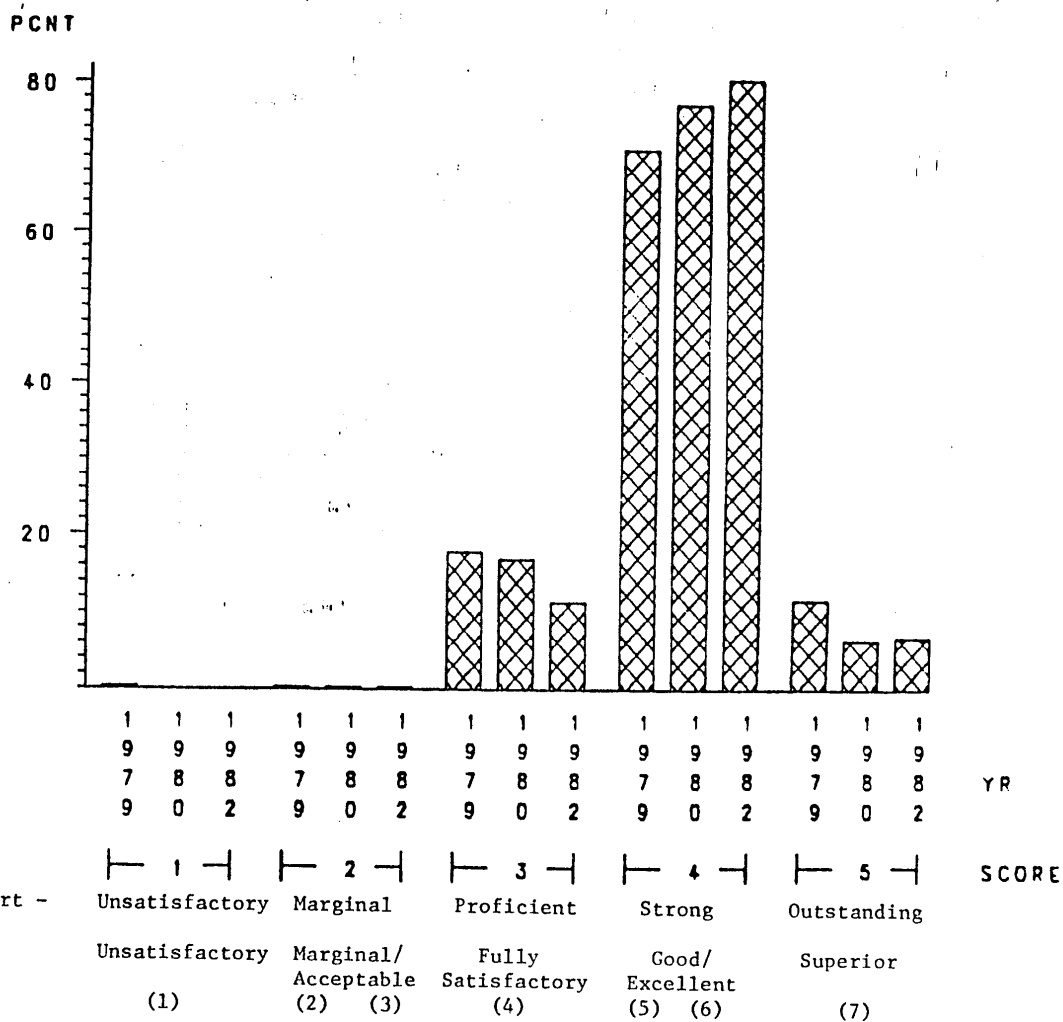
Numbers ⎫
        ⎬ Fitness Report –   Unsatisfactory   Marginal   Proficient   Strong   Outstanding
Compare ⎬
        ⎬ PAR –   Unsatisfactory   Marginal/   Fully   Good/   Superior
With   ⎭   Acceptable   Satisfactory   Excellent
                               (1)   (2)   (3)   (4)   (5)   (6)   (7)
```

# AVERAGE PARS SCORE BY GRADE
## FOR FY 1982